

Development of error criteria for adaptive multi-element polynomial chaos approaches

B. Chouvion^a, E. Sarrouy^{b,*}

^a*LMA, Aix-Marseille Univ, CNRS, UPR 7051, Centrale Marseille, F-13402 Marseille
Cedex 20, France*

^b*LMA, Centrale Marseille, CNRS, UPR 7051, Aix-Marseille Univ, F-13451 Marseille,
France*

Abstract

This paper presents and compares different methodologies to create an adaptive stochastic space partitioning in polynomial chaos applications which use a multi-element approach. To implement adaptive partitioning, Wan and Karniadakis first developed a criterion based on the relative error in local variance. We propose here two different error criteria: one based on the residual error and the other on the local variance discontinuity created by partitioning. The methods are applied to classical differential equations with long-term integration difficulties, including the Kraichnan-Orszag three-mode problem, and to simple linear and nonlinear mechanical systems whose stochastic dynamic responses are investigated. The efficiency and robustness of the approaches are investigated by comparison with Monte-Carlo simulations. For the different examples considered, they show significantly better convergence characteristics than the original error criterion used.

Keywords: Polynomial Chaos, MEgPC, Uncertainty Modelling, Stochastic Galerkin Method

1. Introduction

In various applications, engineers and researchers use complex modelling techniques and simulations to predict behaviour of structures and optimise their design process. However, in order to incorporate more realistic data into their
5 models, it is necessary to take uncertainties into account within the design stage and to predict their influence. Different stochastic or probabilistic procedures are used for modelling these uncertainties. A review on the numerical methods for stochastic prediction can be found for instance in [1], or more recently in [2, 3].

*Corresponding author. Tel.: +33 491 054 390.

Email addresses: benjamin.chouvion@ec-lyon.fr (B. Chouvion),
emmanuelle.sarrouy@centrale-marseille.fr (E. Sarrouy)

10 The most straightforward statistical method based on Monte Carlo (MC) simulation [4] is often used as reference. It relies on the calculation of the direct problem for a large sample. The main drawback is probably that its accuracy relies heavily on the sample size and it is therefore computationally expensive. To confront these difficulties and minimise the sampling size, dedicated meth-
15 ods have been developed such as Latin hypercube [5] or quasi-Monte Carlo SOBOL [6] sampling. On the other hand, non-statistical methods that are not based on very large sampling, have also been developed. For instance, the perturbation methods [7] are based on a Taylor series expansion of the random field around its mean value. However such methods cannot handle accurately
20 uncertainty with high discrepancy because the theory is based on assumed small perturbations. Furthermore it implies that the function derivative around its mean is finite up to a given order, which may not be possible for not-smooth functions.

The Homogeneous Chaos is another branch of non-statistical methods, initiated by Wiener [8] who used Hermite polynomials to model stochastic processes with Gaussian random variables. Ghanem and Spanos then introduced polynomial chaos with a finite element method to model uncertainty in solid mechanics [9]. They showed exponential convergence rate for Gaussian stochastic processes. The non-intrusive stochastic collocation approach [10] uses direct
30 deterministic, but small, samplings at collocation points in the random space and recreates the stochastic response by constructing local polynomial interpolations. In order to deal with discontinuities in the probability or physical space, it was later extended to the multi-element stochastic reduced basis methods [11], the multi-element probabilistic collocation method (MEPCM) [12, 13],
35 to an adaptive hierarchical sparse grid collocation method [14], or the simplex stochastic collocation [15]. The intrusive approach also makes use of polynomials to model uncertainties. The main idea of the method is to expand the random solution spectrally in the polynomial basis and subsequently use a Galerkin projection scheme to transform the original stochastic problem into
40 a set of deterministic algebraic equations with more unknowns. Xiu and Karniadakis [16] introduced the concept of generalised Polynomial Chaos (gPC) using different polynomial bases defined with the correspondence between the probability density functions of random variables and the weight functions of the scalar product that orthogonalises the basis. It is a generalisation of the
45 original Wiener’s Hermite-chaos and can deal with non-Gaussian random inputs efficiently. Soize and Ghanem later proposed a proper mathematical frame to address the chaos representation of vector-valued random variables with arbitrary distributions [17]. More recently, Wan and Karniadakis [18] showed how to create orthogonal polynomials on the fly to extend the procedure to arbitrary
50 random distributions.

Depending on the problem, one can be confronted with quantities whose variations with respect to the random parameters are not continuous. gPC has difficulties for converging for such discontinuous distributions in the stochastic space. Increasing the maximum degree of the chosen basis (p -refinement) may
55 help but it cannot always overcome those convergence difficulties. It also en-

larges the size of the system to solve and makes the solution of the stochastic problem more complex. It is known for instance that gPC fails to converge for long-term integration (see the Kraichnan-Orszag Problem in [19]). The reason is that a certain shape of solution is assumed at the initial time, but this does not necessarily correspond to the real shape or distribution at later times. For long term integration, the method called time-dependent generalised polynomial chaos (TDgPC) [20] recalculates the polynomial basis after a certain time, which helps for convergence.

The regularity or irregularity of the solution with respect to the stochastic space affects the convergence rate of gPC expansion and is usually not known *a priori* in many cases. One of the strategies to tackle this problem is to divide the stochastic space (*h*-refinement) and use a relatively low degree of expansion on each element of the partition, which corresponds to a piecewise polynomial approximation fitting. This has motivated Wan and Karniadakis to propose the multi-element generalised polynomial chaos (MEgPC) method [19]. The space of random inputs is decomposed in small elements. In each element, a new random variable is created and the standard gPC method is applied. Their mesh adaptation scheme to decompose the space of random inputs is based on the relative error in the variance prediction. The partitioning process is performed by dividing each element which does not satisfy the criterion in two equal-size elements.

As far as the authors know, the error criterion proposed in [19] is the only one that was used in MEgPC problems. It was for instance re-used later by the same authors in [18] for solving differential equations, and more recently in the field of robotics by Kewlani and Iagnemma [21], or to predict limit-cycle oscillations of an elastically mounted airfoil [22]. However, as it will be discussed in further details in this manuscript, this error criterion has some inherent drawbacks. Further developments in multi-element polynomial chaos partitioning methodology are therefore needed and these are the main objectives of the present manuscript. Two new error criteria are proposed here to enhance polynomial chaos capabilities for the solution of stochastic differential or algebraic equations.

The manuscript is organised as follows. The mathematical formulation of the generalised Polynomial Chaos (gPC) and its extension to multi-elements are briefly presented in Section 2. Then, the proposed error criteria are described in Sections 3. Numerical results and comparison with reference studies provided in [19] are discussed in details in Section 4.

2. Polynomial Chaos and its extension to Multi-Element approaches

This section briefly describes the principles of Polynomial Chaos expansion and its extension to a multi-element approach when considering the uncertainty propagation in a problem initially defined by a set of deterministic governing equations. It is assumed that uncertainty affects some parameters of these governing equations and that the stochastic model of these *input variables or parameters* is known. The Polynomial Chaos is then applied in order to obtain a

100 stochastic description of some other quantities of interest which will be referred to as *output variables*.

In what follows, two types of governing equations are considered: sets of algebraic governing equations, denoted $H(u) = 0$ and ordinary differential equations (ODE) denoted $\dot{u} = f(u, t)$.

105 2.1. Probability space and notations

Let define $(\Theta, \mathcal{A}, \mathbb{P})$ a probability space with Θ the event space, \mathcal{A} the σ -algebra on Θ and \mathbb{P} a probability measure. The probability density function (pdf) associated to a random variable U is denoted p_U . U expected value is defined as:

$$\mu_U = \mathbb{E}[U] = \int_{\Theta} U(\theta) d\mathbb{P}(\theta) = \int_{\mathbb{R}} u p_U(u) du. \quad (1)$$

110 Finally, in the rest of the paper, a one-dimension stochastic space is considered to simplify the notations and developments. This means that a single random variable is used to introduce uncertainty in the systems. Let denote ξ this random variable whose distribution will be defined later:

$$\xi : \begin{matrix} \Theta & \longrightarrow & \mathbb{R} \\ \theta & \longmapsto & \xi(\theta). \end{matrix} \quad (2)$$

2.2. Polynomial Chaos

115 The three main steps for the implementation of the polynomial chaos are briefly recalled below. For more detailed explanations, the reader is referred to the references cited therein and the book [23].

Polynomial Chaos Expansion. In the polynomial chaos approach, a second-order random process U is expressed as a polynomial series using a set of orthogonal polynomials ψ_k in the variable ξ . In numerical implementation, the sum is usually truncated to a polynomial order (or maximum total degree) p leading to an approximation \hat{U} of the random process U . When the stochastic space dimension is equal to 1, the number of polynomials considered is $N_p + 1$ with $N_p = p$; general formulas for the multi-dimensional case can be found in [16].
125 p -convergence has been proved for any second-order process in the L_2 sense [24]. To simplify the notation, no distinction will be made between U and \hat{U} :

$$U(\theta) = \hat{U}(\xi(\theta)) := \sum_{k=0}^{N_p} u_k \psi_k(\xi(\theta)), \quad (3)$$

where u_k are the expansion (deterministic) coefficients and – in this one dimensional case – ψ_k is the polynomial of degree k . In ODEs, the time dependence is included in the u_k coefficients.

130 *Choice of a polynomial basis.* The choice of the polynomial basis ψ_k used for the expansion of output variables is not obvious. If one considers that the output (unknown) distribution will be close to the input (known) distribution, then an optimal polynomial basis can be chosen according to the input distribution. This is the main idea of the generalised Polynomial Chaos (gPC), based on the
135 Wiener-Askey scheme, as demonstrated by Xiu and Karniadakis [16]. However, if the output variable distribution is very different from the input distribution, there is no obvious reason to make such a choice, except for the sake of simplicity. Recent studies have addressed this issue for time-dependent problems [20]: the random variable ξ used in (3) as well as the polynomial basis ψ_k are regularly
140 updated in order to keep a low degree of expansion.

Expansion coefficients evaluation. The last step is to choose a method to evaluate u_k coefficients. To this end, two families of methods are to be found in the literature: the family of intrusive methods and that of non-intrusive methods.

The intrusive method follows a Galerkin approach: the expressions of the
145 variables expanded on the polynomial basis are re-inserted in the general governing equations of motion; then a scalar product is taken between each of the n_{eq} governing equations and each of the $N_p + 1$ polynomials constructing the p order polynomials basis. This leads to a larger system of equations to solve whose size is $n_{\text{eq}} \times (N_p + 1)$ but does not require any direct evaluation of a
150 realisation of the output variable U .

On the other hand, the non-intrusive methods rely on the vast family of collocation methods; they are generally referred to as Probabilistic Collocation Methods (PCM). Among these methods, is the cubature based method [22]. It uses the orthogonality between the polynomials forming the basis of expansion with respect to a given scalar product denoted $\langle \bullet, \bullet \rangle$. The unknown
155 coefficients u_k are then found with

$$u_k = \frac{\langle U, \psi_k \rangle}{\langle \psi_k, \psi_k \rangle} = \frac{1}{\langle \psi_k, \psi_k \rangle} \int_{\mathbb{R}} U(\xi) \psi_k(\xi) w(\xi) d\xi, \quad (4)$$

where $w(\xi)$ is the weight associated with the scalar product orthogonalising the basis. In this latter expression, the integral is generally evaluated via a quadrature rule. This implies evaluating several realisations of U for determined
160 ξ values [25].

A discussion on the choice of the method depending on the problem under consideration is provided in [23, Sec. 7.4].

2.3. Multi-Element Polynomial Chaos

From what precedes, generalised Polynomial Chaos can be interpreted as a
165 polynomial approximation of a function U depending on ξ . In the case when U has singularities or is discontinuous (see for instance in [26]) or when high polynomial orders are required, a natural solution is to split the range in ξ into a collection of elements and to use a low order polynomial approximation on each element. That is basically what MEgPC and MEPCM rely on.

170 Finally, it is now assumed that the random variable ξ used to introduce uncertainty in the system as well as to expand random processes follows a uniform distribution:

$$\xi \sim \mathcal{U}(-1; 1); \forall x \in [-1, 1], p_\xi(x) = \frac{1}{2} \quad (5)$$

The detailed procedures of multi-element approaches for the uniform distribution case as well as for a more general case can be found in [19, 18] (MEgPC) and [12] (MEPCM).

175 The basic principle of the multi-element methods is to split ξ definition domain $I = [-1, 1]$ in N_e ad-joint elements $I_n = [\xi_n, \xi_{n+1}]$, $n \in \llbracket 1, N_e \rrbracket$, with $\xi_1 = -1$ and $\xi_{N_e+1} = 1$. Over each element I_n , a change of variable is applied:

$$\forall n \in \llbracket 1, N_e \rrbracket, \quad \xi \in I_n \Rightarrow \xi = \alpha_n + \beta_n \zeta_n, \quad \zeta_n \in [-1, 1], \quad (6)$$

180 with $\alpha_n = \frac{\xi_n + \xi_{n+1}}{2}$ the element middle point and $\beta_n = \frac{\xi_{n+1} - \xi_n}{2}$ the element half-length. The new variable ζ_n varies on $[-1, 1]$ as one has $\zeta_n = -1$ for $\xi = \xi_n$, and $\zeta_n = 1$ for $\xi = \xi_{n+1}$. Assuming a uniform distribution of ξ on $[-1, 1]$ implies a uniform distribution of ζ_n also on $[-1, 1]$.

The random process U is then approximated on each element I_n by gPC and is written as:

$$U_n = \sum_{k=0}^{N_{p_n}} \tilde{u}_{k,n} \psi_k(\zeta_n), \quad (7)$$

185 where $\tilde{u}_{k,n}$ is the $(k+1)^{\text{th}}$ coefficient of U expansion on I_n , p_n is the polynomial order used on I_n , and $N_{p_n} + 1$ is the size of the polynomial basis used for the expansion over I_n . The $\tilde{u}_{k,n}$ coefficients can be evaluated using an intrusive (MEgPC) or non-intrusive (MEPCM) approach.

190 In this multi-element approach, there are two main choices to make: the choice of the expansion order p_n over each element (defining N_{p_n}) and the choice of the partition I_n , $n \in \llbracket 1, N_e \rrbracket$. In what follows, as the main focus is on the partition creation, a constant p_n order, $p_n = p$, $n \in \llbracket 1, N_e \rrbracket$ is considered, which means that the polynomial approximation is assumed to have the same order on every element. As an *a priori* optimal partition is almost impossible to determine for the general case, an adaptive partition scheme has to be implemented. Hence, a distinction is made between two important characteristics of such a scheme:

- **the error criterion** : Numerical value or property that each element must satisfy to be kept for the final partition. Different error criteria are presented in Section 3.
- **the partitioning process** : Methodology that defines how elements are divided (or created) in order to converge towards a final partition in which every element satisfies the error criterion. This point is not addressed in this paper and only standard methods are used (see Section 3.3).

205 3. Adaptive error criterion

Two error criteria are proposed and compared to the one initially proposed by Wan and Karniadakis [19] - referred to as criterion I – which is exposed first and considers the variance decay rate. The first alternative, referred to as criterion II, draws a parallel with refinement methods used in finite element
 210 meshing and relies on the variance continuity. The second alternative, referred to as criterion III, proposes a change in the quantity of interest and takes account of the residual error on the initial governing equations.

As the random variable ξ is assumed to have a uniform distribution, ψ_k polynomials will henceforth denote Legendre polynomials [16] with the conventional
 215 notations recalled in Appendix A.

3.1. Error on the variance convergence

As criteria I and II rely on the variance convergence, the expressions of local and global mean and variance in a multi-element frame are developed here.

The local mean μ_{U_n} and variance $\sigma_{U_n}^2$, on a particular element I_n approximated by polynomial chaos (order p , number of polynomials $N_p + 1$), are defined as:

$$\mu_{U_n} = \tilde{u}_{0,n} , \quad (8)$$

$$\sigma_{U_n}^2 = \sum_{k=1}^{N_p} \tilde{u}_{k,n}^2 < \psi_k, \psi_k > . \quad (9)$$

The approximate global mean μ_U and variance σ_U^2 are then given by [19]:

$$\mu_U = \sum_{n=1}^{N_e} \tilde{u}_{0,n} \beta_n , \quad (10)$$

$$\sigma_U^2 = \sum_{n=1}^{N_e} \left(\sigma_{U_n}^2 + (\tilde{u}_{0,n} - \mu_U)^2 \right) \beta_n . \quad (11)$$

Hence, the global mean, or variance, can be seen as the weighted sum of the
 220 mean, or variance respectively, of each element.

3.1.1. Criterion I : use of the local decay rate

The original error criterion proposed by Wan and Karniadakis [19] is based on the variance convergence when either the polynomial order of approximation or the number of elements in the partition increases. It is used for comparison in
 225 the numerical simulations presented in Section 4, and will be recalled thereafter as “criterion I”.

Wan and Karniadakis define the local decay rate of relative error of the PC approximation on the variance in I_n as [19]:

$$\eta_n^I = \frac{\tilde{u}_{N_p,n}^2 < \psi_{N_p}, \psi_{N_p} >}{\sigma_{U_n}^2} = \frac{\tilde{u}_{N_p,n}^2 < \psi_{N_p}, \psi_{N_p} >}{\sum_{k=1}^{N_p} \tilde{u}_{k,n}^2 < \psi_k, \psi_k >}. \quad (12)$$

230 This local decay rate shows the influence of the last term (in degree p) of the polynomial expansion on the local variance. It seems intuitive to think that the approximation is close to its converged limit (and therefore to the exact value) when this extra term does not have a significant influence on the local variance.

The n -th element is then refined when the following condition is satisfied:

$$\beta_n (\eta_n^I)^\alpha > \theta_1, \quad 0 < \alpha < 1, \quad (13)$$

where θ_1 is a constant of accuracy required, and α is usually set at $1/2$ [19, 21].

235 The decay rate of relative error is multiplied by the factor of β_n in Eq. (13) with the idea that an error on the local variance on a small element should contribute only little to the global error. This weighting relaxes the restriction on the accuracy of the local variance for smaller elements.

240 For further use, let us define the error on the n -th element I_n for criterion I as follows:

$$\epsilon_n^I = \beta_n (\eta_n^I)^\alpha, \quad (14)$$

and the global error as $\epsilon^I = \sum_{n=1}^{N_e} \epsilon_n^I$.

The following limitations of the proposed criterion have motivated the authors to offer some new alternatives:

- 245 • It is not relevant to use simple piecewise linear regression ($p = 1$) with this error indicator. With an approximation of degree $p = 1$, the term η_n^I is constant for all elements and equal to 1. This means that the error criterion ϵ_n^I becomes only function of the length of the elements, and no difference is seen between equal-length elements.
- 250 • If the stochastic response function is odd with respect to the random parameter, then the coefficients in $\tilde{u}_{k,n}$ with k even are very small compared to the coefficients with k odd. In the intrusive approach, they are not exactly equal to 0 because the coefficients are all coupled in the equations; whereas they are null if the non-intrusive approach is used. This means that if the maximum order of expansion is even, the error criterion will always be very small and will probably never satisfy Eq. (13). Similar reasoning can be used if the order of expansion is odd for an even function with respect to the random parameter.
- 255 • For particular problems such as the ODE presented in Section 4.3.1, this criterion associated with dichotomy partitioning, as in the reference paper [19], does not seem very efficient as it creates only uniform meshes whatever the order of expansion (see for instance Fig. 9.(a.2)).
- 260 • It may be delicate to find an appropriate θ_1 value because the physical sense of variance convergence is somehow difficult to quantify. The development of another criterion with more meaningful sense would be interesting and this is the meaning of the authors development in Section 3.2.
- 265

3.1.2. Criterion II : use of the variance discontinuity

This section shows a new way to calculate a local error for each of the elements segmenting the stochastic space. It is based on the error in the continuity of a parameter linked to the variance over the stochastic space. It will be used in numerical applications in Section 4 in which it is referred to as “criterion II”.

In the multi-element methods, one wants to approximate $U(\xi)$ by the use of several polynomial functions, each of them defined on ad-joint elements, see Section 2. Let us assume that the function $U(\xi)$ is at least C^1 over $[-1, 1]$ and its derivative with respect to ξ is denoted $U'(\xi)$. Then over each $I_n = [\alpha_n - \beta_n, \alpha_n + \beta_n]$ element, the variance of U_n can be written as

$$\sigma_{U_n}^2 = \int_{-1}^1 (U_n(\zeta_n) - \mu_{U_n})^2 \frac{1}{2} d\zeta_n = \frac{1}{12} (U'(\alpha_n))^2 (2\beta_n)^2 + o(\beta_n^2), \quad (15)$$

and so

$$\lim_{\beta_n \rightarrow 0} \frac{\sigma_{U_n}}{2\beta_n} = g(\alpha_n) \text{ with } g(\xi) := \frac{\sqrt{3}}{6} |U'(\xi)|. \quad (16)$$

Hence, the ratio

$$\hat{g}_n = \frac{\sigma_{U_n}}{2\beta_n} \quad (17)$$

can be considered as a zeroth-order approximation over I_n of the continuous function g defined in Eq. (16).

As g cannot be evaluated in its current definition, it is reconstructed from the known values \hat{g}_n . This reconstruction is similar to what is used in stress smoothing for finite element calculation [27] where predicted stresses show discontinuities at the element edges. The solution $\hat{g}(\xi)$ is the junction of the solutions \hat{g}_n defined on each element I_n . $\hat{g}(\xi)$ shows discontinuities at each element break, similar to the stresses in classical finite element methods. With a smoothing technique, a solution $\tilde{g}(\xi)$ continuous on the whole stochastic space is then built.

This method is somehow similar to the derivative recovery technique developed by Zienkiewicz and Zhu [28, 29, 30] and is illustrated in Fig. 1. Basically, it consists in finding \tilde{g} as a linear regression by part between the different \hat{g}_n , without imposing slope continuity:

- On each element centre α_n , the value of \tilde{g} is taken equal to \hat{g}_n :

$$\tilde{g}(\alpha_n) = \hat{g}_n, \quad n \in \llbracket 1, N_e \rrbracket. \quad (18)$$

Linear segments are used between $\tilde{g}(\alpha_n)$ and $\tilde{g}(\alpha_{n+1})$. This defines \tilde{g} over $[\alpha_1, \alpha_{N_e}]$.

- For the last pieces $[-1, \alpha_1]$ and $[\alpha_{N_e}, 1]$, \tilde{g} is defined as an extrapolation of \tilde{g} over $[\alpha_1, \alpha_2]$ and $[\alpha_{N_e-1}, \alpha_{N_e}]$ respectively. In the case when these extrapolations would induce a negative value for $\tilde{g}(\xi_1)$ or $\tilde{g}(\xi_{N_e+1})$, this value is set to zero, as illustrated in Fig. 1.

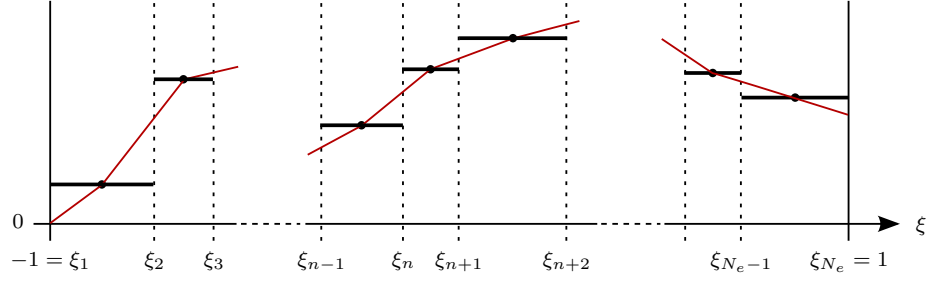


Figure 1: Smoothing technique to create \tilde{g} from \hat{g} . Thick (black) lines: \hat{g} ; thin (red) lines: \tilde{g} .

The error on an element is then evaluated using the absolute value of the difference between \tilde{g} and \hat{g}_n . This difference is adimensioned by dividing it by $\hat{g}_{av.}$ defined as:

$$\hat{g}_{av.} = \sum_{n=1}^{N_e} \beta_n \hat{g}_n = \frac{1}{2} \sum_{n=1}^{N_e} \sigma_{U_n}. \quad (19)$$

This leads to:

$$\epsilon_n^{II} = \int_{\xi_n}^{\xi_{n+1}} \frac{|\tilde{g}(\xi) - \hat{g}_n|}{\hat{g}_{av.}} d\xi = \beta_n \int_{-1}^1 \frac{|\tilde{g}(\alpha_n + \beta_n \zeta_n) - \hat{g}_n|}{\hat{g}_{av.}} d\zeta_n = \beta_n \eta_n^{II}. \quad (20)$$

A detailed expression of η_n^{II} is provided in [Appendix B](#).

The global error is equal to

$$\epsilon^{II} = \sum_{n=1}^{N_e} \epsilon_n^{II} = \int_{-1}^1 \frac{|\tilde{g}(\xi) - \hat{g}(\xi)|}{\hat{g}_{av.}} d\xi. \quad (21)$$

The random element is refined when the following condition is satisfied:

$$\beta_n \eta_n^{II} > \theta_2, \quad (22)$$

where θ_2 is a given constant of accuracy.

It may be difficult to give a physical meaning to the value of acceptable threshold θ_2 . A possible good partitioning strategy may therefore be not to use Eq. (22) criterion but to refine the element(s) with the worst error indicator Eq. (20) until reaching a maximum number of elements, or a maximum computation time.

3.2. Criterion III : residual error

Once the approximation by PC expansion is performed on a particular element of the stochastic space, one can evaluate its error from the real solution by looking at the residual error this approximation creates in the governing system of equations.

When the governing equations are algebraic, the definition of the residue (denoted R) is straightforward:

$$R(U; \xi) = H(U; \xi), \quad (23)$$

where U is defined from the PC decomposition.

320 In the case where the governing equations are a set of ODEs, the calculation of the residue depends on the methods – intrusive or non-intrusive – used to compute the coefficients of the expansion. The basic idea though is to verify that the relationship $\dot{U} = f(U, t)$ is satisfied for some t values. Hence, the residue is defined as follows

$$R(U; \xi) = \dot{U} - f(U, t; \xi). \quad (24)$$

325 The computation of \dot{U} from U PC expansion is detailed in [Appendix C](#). With the above definition of the residue for a given value of ξ , a global error can be evaluated as:

$$\epsilon^{\text{III}} = \int_{-1}^1 |R(U; \xi)| \frac{1}{2} d\xi. \quad (25)$$

This global error can be split into a sum of local errors over each element:

$$\epsilon^{\text{III}} = \sum_{n=1}^{N_e} \int_{\xi_n}^{\xi_{n+1}} |R(U(\xi); \xi)| \frac{1}{2} d\xi = \sum_{n=1}^{N_e} \epsilon_n^{\text{III}}, \quad (26)$$

with the local error on the n -th element being defined as

$$\epsilon_n^{\text{III}} = \beta_n \eta_n^{\text{III}}, \quad \eta_n^{\text{III}} = \int_{-1}^1 |R(U_n(\zeta_n); \alpha_n + \beta_n \zeta_n)| \frac{1}{2} d\zeta_n. \quad (27)$$

330 The random element is refined when the following condition is satisfied:

$$\beta_n \eta_n^{\text{III}} > \theta_3, \quad (28)$$

where θ_3 is a given constant of accuracy.

This error indicator is referred to as “criterion III” in the remaining of the manuscript, and will be compared to criteria I and II in [Section 4](#).

335 The numerical value of θ_3 to use will depend on the set of governing equations H : it can indeed be correlated to the termination tolerance on the function value usually employed in numerical solvers.

3.3. Remark on the partitioning process

340 The most common and simplest partitioning process consists in splitting each element needing refinement in half and generating two equal-size elements. This is a simple dichotomy procedure and it was employed for instance in [\[19\]](#). If the global method needs, or can take advantage of, the result contained in the mother element, a simple interpolation, as well as the associated particular change of variable, is achieved (see details in [\[19\]](#)). By this means, the random

field is mapped from one mesh of elements to a new mesh of elements and this
 345 acts as initialisation for the new system solution.

The dichotomy partitioning method can be applied in different ways. The
 first one, used in [19], consists in splitting simultaneously all the elements that
 do not satisfy the given error criterion with threshold θ_i . This iterative process
 ends when there are no elements left to split, and therefore all the remaining
 350 elements satisfy the given error criterion. As all of the elements are studied
 independently of each other, this process can be easily performed in parallel
 computing. Another method is to split only the element with the worst error
 indicator, and repeat the process until a maximum computation time, a maxi-
 mum number of elements or a minimum element size has been reached. For a
 355 given computation time or a given number of elements, the mesh and solutions
 obtained with the two aforementioned methods are different.

Developing new partitioning processes is beyond the scope of this study.
 Suggestions for possible development include reusing strategies that are widely
 employed in classical finite element analysis (see for instance [27]), or invoking
 360 some of the principles of continuation [26].

4. Numerical applications

The different error criteria proposed in Section 3 are here illustrated with
 benchmark examples, and their convergence characteristics compared. First,
 cases when U is a random variable or random vector (not time-dependent) are
 365 analysed, those are typical linear and nonlinear mechanical oscillators. Then,
 cases when U is a random process indexed by time, solution of a set of differential
 equations, are studied. In all of the applications, the one-dimensional random
 variable ξ is assumed to follow a uniform distribution on $[-1, 1]$ ($\xi \sim U(-1, 1)$).
 The reference results are those obtained with a Monte Carlo simulation (MC)
 370 and a sampling of 10^6 ξ -values.

4.1. Wasserstein distance

As the post-processing of results usually involves rebuilding the pdf, the
 quality of the approximation cannot be solely assessed by its mean and variance
 convergences. Therefore, the Wasserstein distance [31, Chap. 6] to the reference
 375 simulation is hereby also used. This distance provides a means to compare the
 quality of the approximation considering simultaneously the convergence on the
 mean, the variance and the distribution law.

The Wasserstein distance (Wd) of order p between two probability measures
 P_1 and P_2 having a second order moment is defined as [31]:

$$\text{Wd}_p(P_1, P_2) = \inf \left\{ \mathbb{E}[(X_1 - X_2)^p]^{1/p} : \mathcal{L}(X_1) = P_1, \mathcal{L}(X_2) = P_2 \right\}, \quad (29)$$

380 where $\mathcal{L}(X)$ denotes the law of X . The Wasserstein distance of order 2 will be
 used thereafter. In this case, it can be evaluated by:

$$\text{Wd}_2(P_1, P_2) = \left(\int_0^1 (F_{X_1}^{-1}(t) - F_{X_2}^{-1}(t))^2 dt \right)^{1/2}, \quad (30)$$

where F_X denotes the cumulative distribution function associated to X .

The distance between the MEgPC approximation and the reference sample U_{ref} can then be evaluated using Eq. (30). Finally, it is interesting to render this distance dimensionless with respect to the order of magnitude of U [32]. In this way, let us denote Wd the Wasserstein distance of order 2 divided by the standard deviation evaluated using the reference sample:

$$\text{Wd}(P_U, P_{U_{\text{ref}}}) = \left(\frac{\int_0^1 (F_U^{-1}(t) - F_{U_{\text{ref}}}^{-1}(t))^2 dt}{\sigma_{U_{\text{ref}}}^2} \right)^{1/2}. \quad (31)$$

To illustrate this distance, several numerical examples are provided in [Appendix D](#).

390 4.2. Random vectors

4.2.1. Simple linear mechanical system

The first comparisons are performed by studying a simple linear dynamical system with a single degree of freedom q (abbreviated SDof in what follows). Uncertainty is assumed in its undamped natural resonant frequency ω_0 such that $\omega_0^2(\xi) = \overline{\omega_0^2}(1 + 0.3\xi)$. Under unitary harmonic forcing, its equation of motion writes:

$$\ddot{q} + 2\eta\omega_0\dot{q} + \omega_0^2q = \cos(\omega t), \quad (32)$$

with η the damping factor. Let set $\eta = 0.05$ and $\overline{\omega_0^2} = (2\pi)^2$ for the following numerical applications. The steady state response of such a system has the form:

$$q(t) = u \cos(\omega t) + v \sin(\omega t). \quad (33)$$

400 One here focuses on the cosine component u which is evaluated using:

$$u = \text{Re} \left(\frac{1}{\omega_0^2 - \omega^2 + 2j\eta\omega_0\omega} \right) \quad (34)$$

where $\text{Re}(\bullet)$ denotes the real part and $j^2 = -1$. The coefficients of the polynomial chaos expansions over each element are evaluated using a non-intrusive method (see Sec. 2.2). Hence, u will be evaluated repeatedly at the quadrature nodes using Eq. (34).

405 First, the influence of the polynomial degree p used for the expansion is studied, and is illustrated in Fig. 2. The results obtained when using criterion I and criterion III for given thresholds values θ_1 and θ_3 are compared. The quality of the approximation is measured by the dimensionless Wasserstein distance Wd defined in Eq. (31). Thresholds values $\theta_1 = 5 \times 10^{-3}$ and $\theta_3 = 2 \times 10^{-4}$ are chosen so as to obtain similar Wasserstein distances for both criteria. The excitation frequency range ω considered is such that $\omega/(2\pi) \in [0.8; 1.2]$, which is centred around the nominal resonant frequency and large enough to encompass the ω_0 range. The integral of the residue required for criterion III evaluation (see Eq. (27)) is computed using $2p+1$ Gauss integration points. Figure 2 shows

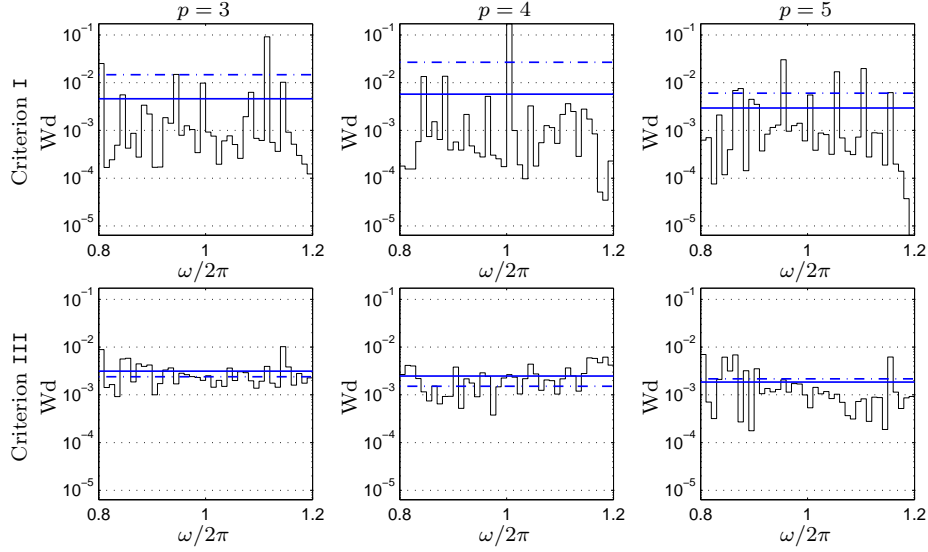


Figure 2: SDof - Evolution of Wasserstein distance with the approximation degree. First row: criterion I results; second row: criterion III results. From left to right, polynomial degree is equal to 3, 4 and 5. Stairstep curves represent Wd with respect to ω while straight lines represent the mean (solid line) and the standard deviation (dashed dotted line) of Wd over the whole ω range.

the Wasserstein distance, as well as its mean and standard deviation over the whole ω -range, obtained using either criterion I or criterion III for p -values equal to 3, 4, or 5, as a function of ω . It can be seen that there are large discrepancies on this mean for criterion I when stepping up from one degree to another while it globally decreases when using criterion III. Moreover the variation of the Wasserstein distance from one ω -value to another is larger when applying criterion I than criterion III. Finally, Fig. 2 also indicates that a poor quality approximation is obtained for $\omega = 2\pi$ when using a $p = 4$ expansion together with criterion I.

The corresponding results, when $\omega = 2\pi$, are drawn in Fig. 3. One can see that the shape of U -pdf is suddenly badly approximated for criterion I and $p = 4$ whereas results for $p = 3$ and $p = 5$ are much more accurate. Here appears the issue mentioned at the end of Section 3.1.1 as the target function comes close to an odd function. Hence, even though the θ_1 value was tuned so as to obtain accurate results at the beginning of the ω -range, the quality of the results over the whole range largely varies. Furthermore, this particular inaccuracy occurs at resonance, a point of major interest in mechanical systems. In this way, Criterion III seems more robust as the results are globally less spread out, considering different ω -values or different degrees.

Figure 4 illustrates the variation of the Wasserstein distance with respect to threshold values θ_1 and θ_3 for a given degree $p = 4$. For each criterion, one considers 5 values of ω around 2π , and 31 values of θ_1 and θ_3 equally spaced on

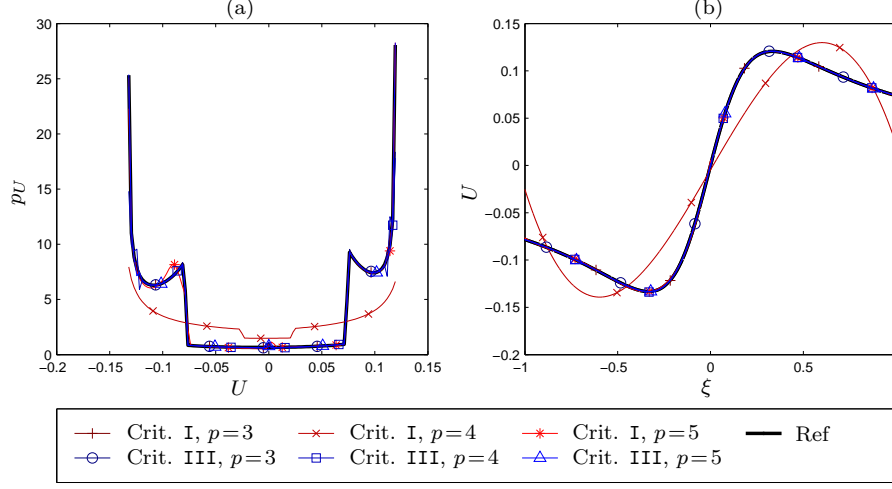


Figure 3: SDof - Comparison of results for different degrees when $\omega = 2\pi$. (a) U -pdf, denoted p_U ; and (b) U as a function of ξ .

a logarithmic scale. The ranges chosen are $[10^{-8.1}, 10^0]$ for θ_1 and $[10^{-12}, 10^{-1}]$ for θ_3 . These ranges provide approximately equal limits (10^{-1} and 10^{-9}) for the Wasserstein distance as depicted in Fig. 4.(a,b) with solid lines. The first observation is that, when the thresholds decreases, the Wasserstein distance for criterion III decreases smoothly and almost in an affine relationship in a log-log scale, while that for criterion I decreases step by step and spreading out more.

The second information depicted is relative to the number of elements produced by each method. Figure 4.(c,d) shows this number of elements as a function of the Wasserstein distance obtained. It then appears that criterion I produces more elements than criterion III to reach an equivalent value of the Wasserstein distance. Hence, the numerical cost is higher, as illustrated in Fig. 5. As cpu times are very small and then difficult to measure and compare, the numerical cost is evaluated through the number of evaluations of Eq. (34) used to compute polynomial chaos expansions. This number of evaluations is represented along Wd in Fig. 5.(a) and the ratio between the number of evaluations required when using criterion III and the number of evaluations needed when using criterion I is displayed on Fig. 5.(b). As data were not accessible for the same abscissa (Wd value) for both criteria, the number of evaluations for each criterion is averaged per Wd decade to plot Fig. 5.(b).

Figures 4 and 5 indicate that it may be hard to set a proper value for θ_1 as the quality of the subsequent approximation varies irregularly with the threshold and that criterion I generally requires more numerical resources than criterion III.

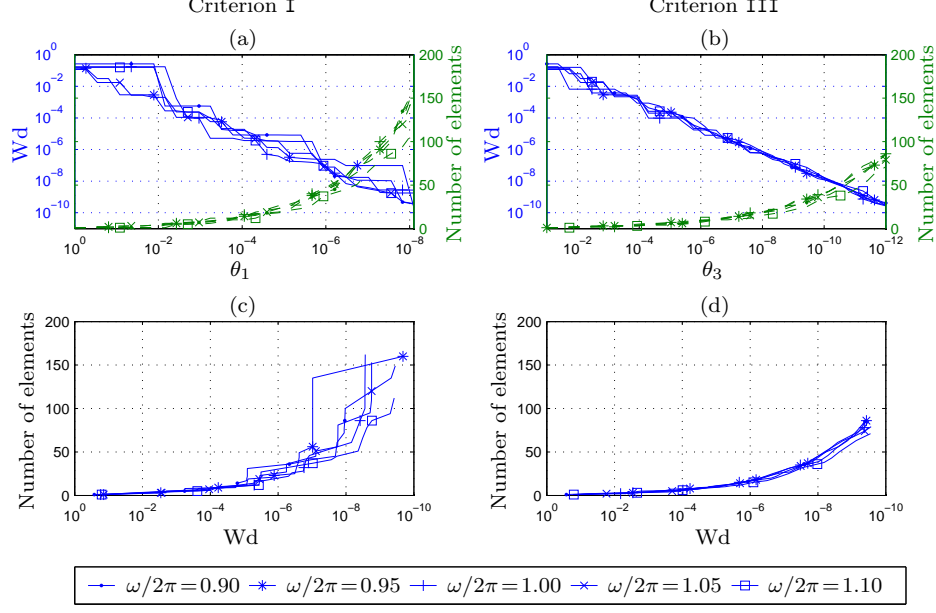


Figure 4: SDof - Evolution of Wasserstein distance (Wd) and number of elements for several values of ω . Left column refers to criterion I and second column to criterion III. The first row provides the evolutions of Wd (solid lines) and the number of elements (dashed lines) with respect to error thresholds θ_1 and θ_3 , while the second row illustrates the evolution of the number of elements with respect to Wd.

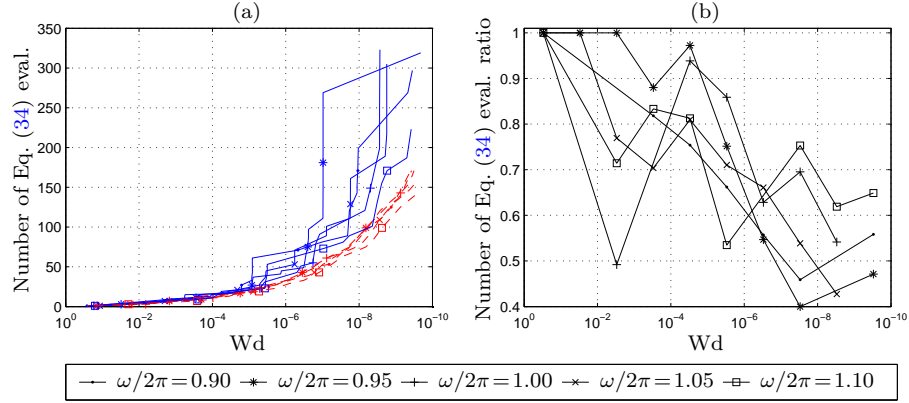


Figure 5: SDof - Numerical cost (number of Eq. (34) evaluations) as a function of the Wasserstein distance (Wd) for several values of ω . (a): (blue) solid lines represent criterion I results; (red) dashed lines represent criterion III results. (b): Ratio of the number of evaluations used by criterion III over the number used by criterion I.

4.2.2. Duffing oscillator

In this section, an uncertain Duffing oscillator is considered as a possible example of the outcome of nonlinear equations derived from a structural mechanics problem.

The governing equation of motion for such an oscillator is nonlinear with a cubic term in stiffness. Under a harmonic excitation, it can be written as:

$$m\ddot{q} + c\dot{q} + kq + k_{nl}q^3 = F \cos(\omega t), \quad (35)$$

where q is the displacement, m the mass, k the stiffness, c the damping coefficient, k_{nl} the purely nonlinear stiffness, F the applied force magnitude, ω the excitation frequency, and t the time.

Looking at the periodic forced response, and using the harmonic balance method [33] with a single term in the Fourier expansion, such as

$$q = a \cos(\omega t) + b \sin(\omega t),$$

Eq. (35) gives the following system, not time-dependent, of nonlinear equations:

$$\begin{cases} (\frac{k}{m} - \omega^2)a + \frac{c}{m}\omega b + \frac{3}{4}\frac{k_{nl}}{m}a(a^2 + b^2) - \frac{F}{m} = 0, \\ (\frac{k}{m} - \omega^2)b - \frac{c}{m}\omega a + \frac{3}{4}\frac{k_{nl}}{m}b(a^2 + b^2) = 0. \end{cases} \quad (36)$$

Uncertainty is assumed in the nonlinear stiffness parameter k_{nl} , and it is defined as $k_{nl} = \bar{k}_{nl}(1 + \alpha_D \xi)$, where \bar{k}_{nl} is its mean and α_D is a coefficient of dispersion.

The unknown coefficients a and b are expanded in the polynomial chaos using Legendre polynomials (see Appendix A) up to a degree p . An intrusive method is used, as explained in Section 2.2, and applied to the system (36), with an order of approximation of $p = 2$. It gives a new system of 6 coupled equations.

Figure 6 shows the nominal displacement amplitude (defined as $\sqrt{a^2 + b^2}$ with $\xi = 0$) as a function of the excitation frequency. The existence of an uncertainty in the nonlinear stiffness implies a set of possible responses. The extreme responses obtained when $\xi = -1$ and $\xi = 1$ are also illustrated. The dispersion of these responses is shown at a given frequency before the tilting of the resonance curve to avoid difficulties in the case one moves forward in the frequency range and reaches a point when either 1 or 3 solutions exist. Darker regions along the dispersion line mean higher probability. One can see that even though the initial uncertainty is taken as uniform, the dispersion is not and results are compacted towards smaller amplitudes.

The numerical solution of the nonlinear system with variables a and b expanded onto the polynomial basis is carried out by a standard nonlinear iterative solver. The error criterion is calculated on every elements. If an element needs refinement, the same kind of mapping as that presented in Section 3.3 is employed to initialise the solution from the mother-element to the sub-elements. Then, a new solution is calculated on the just-formed sub-elements.

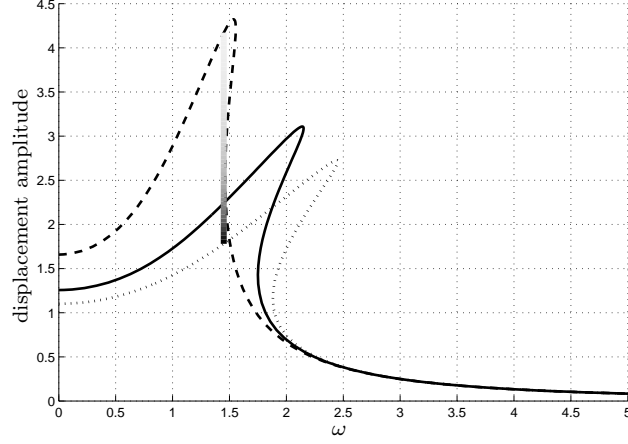


Figure 6: Duffing oscillator - Displacement amplitude as a function of the excitation frequency for different values of ξ . $\xi = -1$: bold dashed line, $\xi = 0$: solid line (nominal value), and $\xi = 1$: light dashed line. The dispersion is shown at the particular frequency of $\omega = 1.45$. Numerical values used are $k = 1$, $m = 1$, $c = 0.3$, $\bar{k}_{nl} = 0.5$, $\alpha_D = 0.8$, and $F = 2$.

Comparisons on convergence characteristics of the different error criteria are illustrated in Figs. 7 and 8. In the dichotomy partitioning process employed, it has been chosen not to split simultaneously all of the elements that do not satisfy the error criterion but to split at each iteration only the element with the worst error. Values of θ_i are therefore here not relevant. In the results labelled “Optimal mesh”, the partitioning process used is different. In this case, a numerical optimisation algorithm looks for the partition for a given number of elements that minimises the global residual error ϵ^{III} (Eq. (25)). The idea behind this simulation is to show the convergence of the multi-element approach when the global residual error is minimised.

Figure 7 illustrates results for coefficient a in Eq. (36) but similar results can be obtained for coefficient b . It gives the approximation error for different indicators as a function of the number of elements created in the partition. It shows that for a given number of elements the proposed error criteria (II and III) give generally better results than the original one. Those are more efficient than ϵ_n^I because dividing elements referring to those errors shows better convergence characteristics. The second row in Fig. 7 illustrates the ratios of the errors with the ones obtained from criterion I and shows that the proposed criteria may offer an increase in quality up to two orders of magnitude.

Figure 8 illustrates the global errors for the different cases. As it could be expected, it shows that faster convergence is obtained on ϵ^i when partitioning using criterion i is performed (i being I, II, or III). The “optimal” mesh considering criterion III gives indeed the best results on ϵ^{III} . Focusing on a particular criterion also tends to decrease the other global errors. This is not

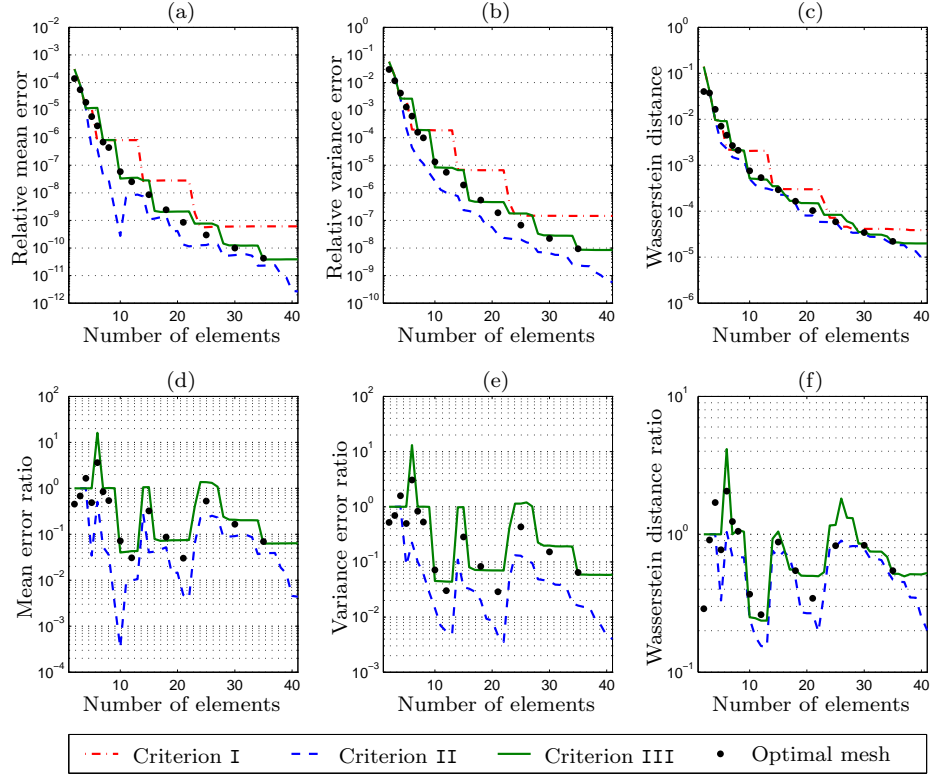


Figure 7: Duffing oscillator - (a) Relative error on the mean, (b) relative error on the variance, and (c) Wasserstein distance, as a function of the number of elements in the partition. (d-f) ratios with results from criterion I for the same error indicators. Numerical values used are those given in Fig. 6.

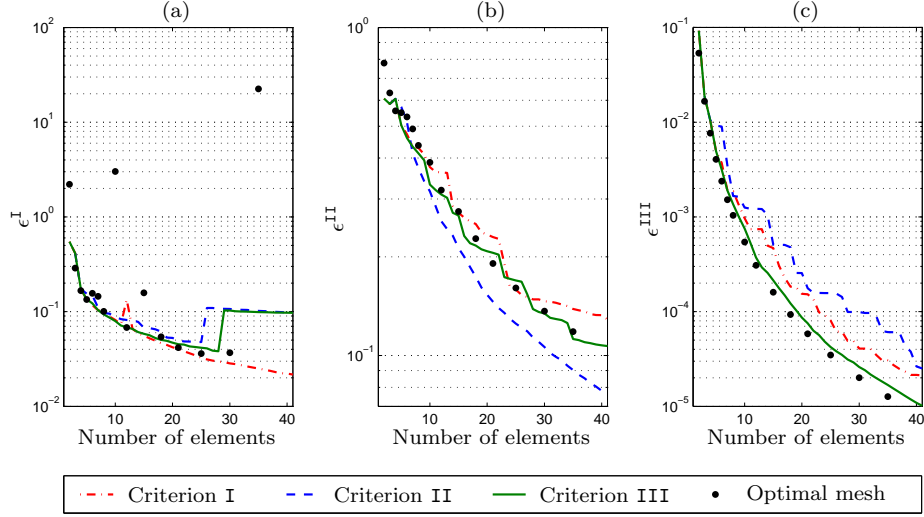


Figure 8: Duffing oscillator - Global error for each criterion, (a) ϵ^I , (b) ϵ^{II} , and (c) ϵ^{III} , as a function of the number of elements in the partition. Numerical values used are those given in Fig. 6.

that clear concerning ϵ^I . It is found for instance that the “optimal” mesh (with respect to minimising ϵ^{III}) may display a partition with more elements (around 10 or 35 elements) but with a worse ϵ^I -value. This shows some kind of lack of robustness in ϵ^I . Furthermore, even when results obtained with criteria II and III show an increase in the global error ϵ^I , they still give mean error, variance error, and Wasserstein distance that decrease as shown in Fig. 7.

4.3. Random processes indexed by time

4.3.1. One-dimensional ODE

In this example, let us focus on the performance of the different error criteria for the following ordinary differential equation (ODE) studied with the original gPC in [16] and the standard MEgPC partitioning in [19]:

$$\frac{du}{dt} = -\xi u, \quad u(0, \xi) = u_{t=0}, \quad (37)$$

The exact analytical solution from which MC simulations are performed is given in Eq. (38). This expression also acts as reference for relative error calculations.

$$u(t, \xi) = u_{t=0} e^{-\xi t}. \quad (38)$$

A similar integration algorithm with adaptive meshing as that explained in [19] is here employed. The time increment used is $\delta t = 0.03$ s for the adaptive scheme. It means that the error criterion evaluation, possible re-meshing, and re-initialisation of initial conditions are performed every 0.03 s. From a given

time t to $t + \delta t$, the integration is performed with a 4th order Runge-Kutta algorithm.

Because it is only possible to calculate criterion II when $N_e \geq 2$, the stochastic space is split in half after the first time increment for the 3 error criteria. Thereafter, the elements that do not satisfy the threshold value at the current time step are split in two equal-length elements, the mapping of the initial conditions is performed, new approximations are calculated and these serve as new initial conditions for the next time step.

First, threshold values θ_i ($i = 1, 2$, or 3) are chosen such that the total number of ODE evaluations at the final time $t = 6$ s is similar for the three different criteria. The number of ODE evaluations for a certain period of time is defined as the number of elements in the partition multiplied by the number of time increments included within this period. For instance a partition with 4 elements used during $10 \delta t$ costs 40 ODE evaluations. This number of ODE evaluations is a direct way to measure the cpu time or complexity of the scheme. Indeed, the calculations of the different error criteria are fast to compute and it is possible to consider that their evaluation time is negligible compared to the cpu time required for the ODE solution. The simulations illustrated in Fig. 9 and Fig. 10 are made with an approximation degree $p = 3$, and threshold values $\theta_1 = 10^{-4}$, $\theta_2 = 8 \times 10^{-3}$ and $\theta_3 = 10^{-6}$, which relate to a total of 4393, 4386, and 4372 ODE evaluations respectively to complete the time integration up to $t = 6$ s. With these simulations, one can compare the behaviour (Fig. 9) and performance (Fig. 10) of the three criteria considering a similar global numerical cost.

Figure 9 illustrates the adaptive mesh created by the different error criteria. It shows that for criterion I, the number of elements is always equal to $N_e = 2^m$ (m integer) and the mesh completely uniform. But for the other criteria, additional splitting is performed towards the left boundary ($\xi = -1$), see Fig. 9.(b.2) and Fig. 9.(c.2), because the exponential curvature of the response (Eq. (38)) is more difficult to approximate polynomially towards this end. One then gets a sensible mesh with element sizes that progressively increase from left to right. It is worth noting that even though the number of elements at $t = 6$ s is greater with the mesh obtained with criterion III than that obtained with criterion I or II, the number of ODE evaluations performed is actually a little bit smaller.

The convergence from adaptive partitioning as a function of time for criterion I has already been shown in [19]. In Figs. 10 and 11, the convergence of the proposed error criteria is compared with that of the original criterion. Figure 10 illustrates the ratio of absolute error obtained with criteria II and III over the absolute error obtained with criterion I, as a function of time with adaptive partitioning. Even though the number of total ODE evaluations is similar for $t = 6$ s, the proposed error criteria offer much better convergence characteristics (smaller errors illustrated by ratios of absolute error smaller than 1) for the mean, the variance and the Wasserstein distance. In Fig. 11, one focuses on the MEgPC approximation state at $t = 6$ s. The θ_i values triggering the partitioning are varied so that different meshes are created along the integration. Smaller θ_i implies more elements and a more accurate approximation, whereas

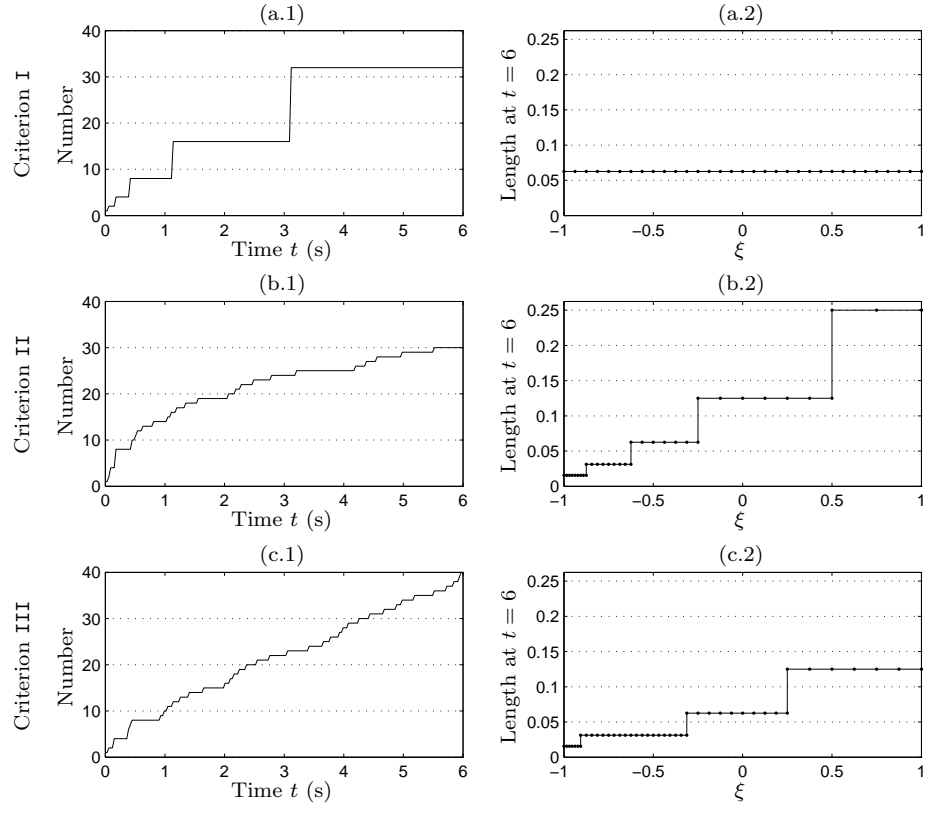


Figure 9: 1D ODE; simulations with a similar final numerical cost (total number of ODE evaluations over the time range) - First column: number of elements in the partition during integration as a function of time t ; second column: adaptive mesh (elements lengths) at $t = 6$ s. Each row contains the results obtained for a particular criterion.

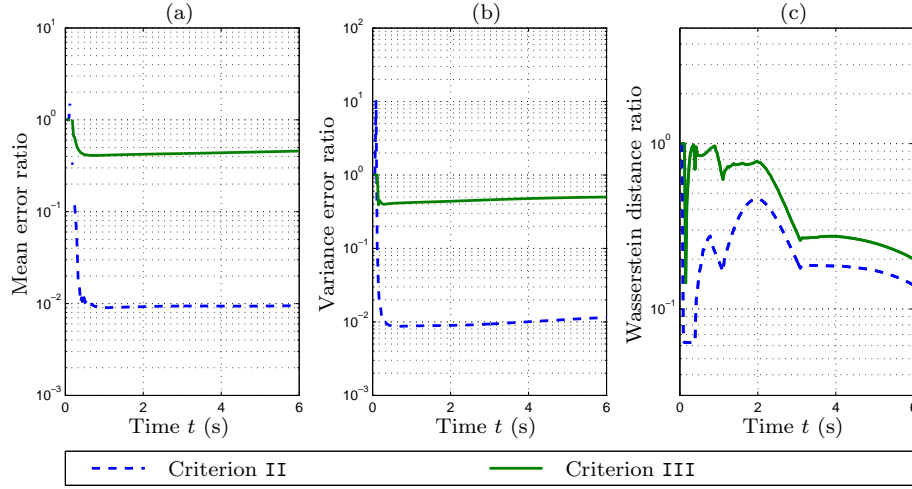


Figure 10: 1D ODE; with similar final numerical cost (total number of ODE evaluations over the time range) - Ratio of absolute error obtained with criteria II and III over absolute error obtained with criteria I, as a function of time during the integration scheme. (a) Error on the mean value, (b) the variance value, and (c) Wasserstein distance.

larger θ_i does not trigger much the partitioning process and implies a rougher approximation. The relative error on the mean value, the variance value, and the Wasserstein distance are illustrated as a function of the number of ODE evaluations (which is itself function of the number of elements in the mesh – or of the value of θ_i). A given number of ODE evaluations directly relates to a given cpu time of computation. For a given computation time, the proposed error criteria, and especially criterion II, show faster convergence than criterion I. It is here obtained better accuracy in fewer ODE evaluations.

Results were shown for an order of expansion $p = 3$ but similar trends and conclusion can be obtained with a higher degree of approximation. However, this simple example is particular as criterion I does not give non-uniform meshes. In the next section, the same methodologies and error criteria are applied to a more complex system of differential equations.

4.3.2. The Kraichnan–Orszag three-mode problem

The Kraichnan–Orszag three mode problem, first introduced in [34], with uncertainty in its initial conditions, is known to fail in a short time when employing gPC. It has been intensively studied to show the loss of accuracy of gPC expansion for problems involving long time integration. This problem is highly sensitive to the initial conditions. It therefore represents a challenging benchmark. Different methods have been applied to this problem such as adaptive MEgPC, TDgPC or adaptive stochastic collocation [19, 18, 12, 20, 14, 35].

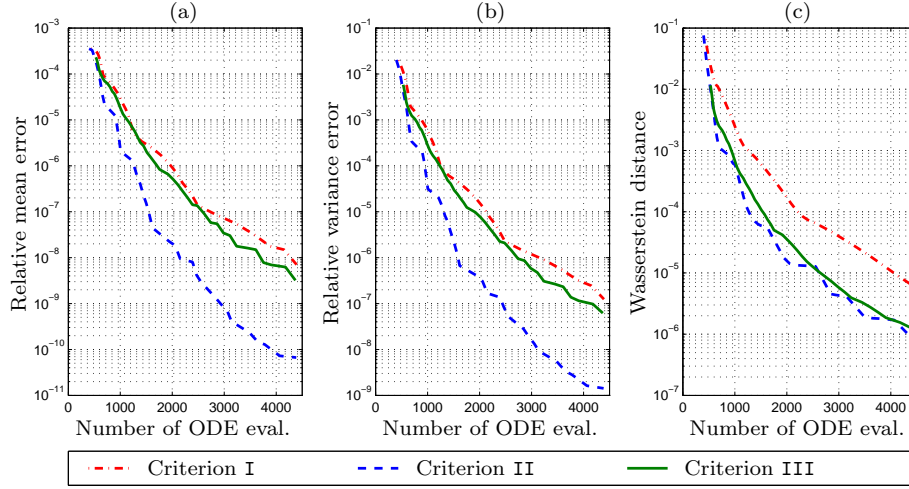


Figure 11: 1D ODE - (a) Relative error on the mean, (b) relative error on the variance, and (c) Wasserstein distance at $t = 6$ s, as a function of the number of ODE evaluations, which is proportional to the numerical cost. The error threshold θ_1 varies between 0.1 and 10^{-4} , θ_2 between 1 and 8×10^{-3} , and θ_3 between 0.1 and 10^{-6} . The order of approximation is $p = 3$.

The problem is defined with the system of ordinary differential equations:

$$\begin{cases} \dot{y}_1 = y_1 y_3, \\ \dot{y}_2 = -y_2 y_3, \\ \dot{y}_3 = -y_1^2 + y_2^2, \end{cases} \quad (39)$$

subject to initial conditions

$$y_1(0) = 1, \quad y_2(0) = 0.1\xi, \quad y_3(0) = 0, \quad (40)$$

The introduction of randomness in the initial conditions implies that the stochastic solution can be either continuous or discontinuous [19].

Wan and Karniadakis [19] showed that MEgPC approximation converges towards MC simulation as θ_1 decreases. In this section the results obtained with MEgPC and with the proposed error criteria are compared with those obtained with criterion I. The order of polynomial approximation is $p = 3$. As in the previous example, the integration is performed with a 4th order Runge-Kutta algorithm. The time increment in the adaptive scheme is 0.03 s. It has been imposed to split the stochastic space in 3 equal-lengths elements after the first iteration because criterion II always needs $N_e \geq 2$, and furthermore, here, it must actually be set at least equal to 3 to avoid problems with criteria I and II due to the parity characteristics of the function. The error thresholds used are $\theta_1 = 8 \times 10^{-3}$, $\theta_2 = 7 \times 10^{-2}$ and $\theta_3 = 1.1 \times 10^{-3}$, which creates partitions with 44, 40 and 40 elements respectively at $t = 30$ s. The error threshold values were chosen such that the number of elements in the mesh at $t = 30$ s is similar for each simulation.

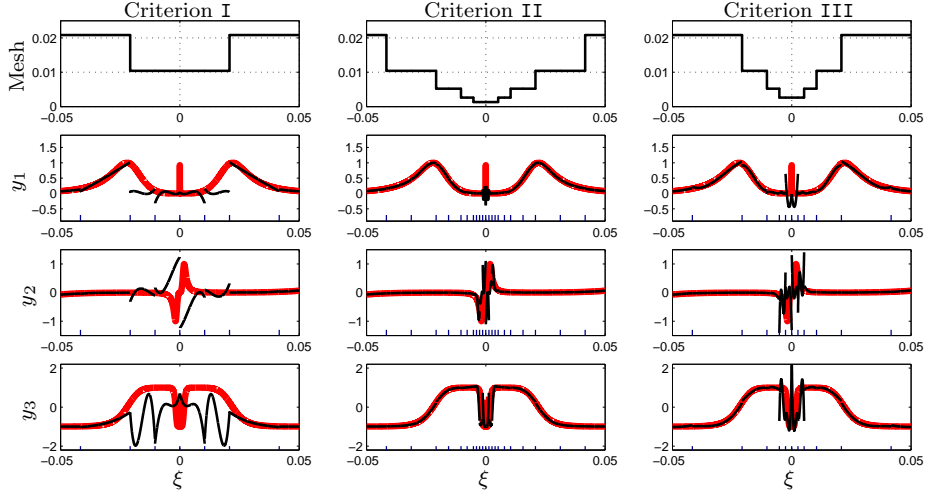


Figure 12: Kraichnan–Orszag problem - Focus around the discontinuity: length of elements in the adaptive mesh (1st row), y_1 -values (2nd row), y_2 -values (3rd row) and y_3 -values (4th row) obtained with the 3 different error criteria at $t = 30$ s. The x-axis is ξ on the interval $[-0.05, 0.05]$. MEGPC approximations (thin black lines) are compared to MC simulations (red thick lines).

The first row of graphs in Figure 12 illustrates the mesh obtained at $t = 30$ s around the discontinuity (at $\xi = 0$) for the different criteria. The results in Fig. 12 show that Criteria II and III produce a finer mesh around the discontinuity. In that way, the area around the discontinuity is better captured. This is illustrated in the following rows of graphs in Figure 12. Those show the random variable y_1 , y_2 , and y_3 values as a function of ξ , compared to the MC simulation. It is not shown in the figure but for the rest of the ξ -range, the 3 methods behave perfectly well. It appears clearly that the meshes obtained with criteria II and III are more efficient than that obtained with criterion I as the approximations around the discontinuity are of better quality. Criterion II behaves particularly well because more partition refinement occurs early in time during integration (similar to what is illustrated in Fig. 9 for the one-dimensional ODE case) and therefore it does carry less inaccuracy along the integration. This makes it possible to keep a relatively good approximation around the discontinuity, whereas for instance with criterion I, the refinement appears too late in time and the approximation obtained after a certain time (for instance at $t = 30$ s) is locally not accurate at all because some inaccuracy was carried out and amplified during the integration scheme.

5. Conclusions

In this paper, different strategies for stochastic space partitioning are proposed. The authors offer alternatives to the initial error criterion suggested by Wan and Karniadakis [19] and illustrate them on several examples. The main

focus was to study relative convergence characteristics: the proposed method-
645 ologies prove to be generally more efficient on the examples tested, and to converge faster in terms of computation time and number of elements, especially for criterion II. It is worth noting that the different error criteria introduced in this manuscript were illustrated on MEgPC problems but could also be adapted without too much difficulty for the adaptive meshing in MEPCM [12].

650 For the initial error criterion suggested by Wan and Karniadakis, it may be difficult to chose the parity of the polynomial approximation. If the output function is odd or even, a wrong choice of polynomial order may end up with not triggering the partitioning process at all. The same kind of problem exists with criterion II and 2 initial elements. However, criterion III is itself really
655 robust and can be applied without any *a priori* knowledge of the regularity or parity characteristic of the uncertain output function.

Only one-dimensional stochastic cases have been considered in this paper. The simple approach consisting in smoothening a piecewise constant function proposed in Section 3.1.2 will need to be extended to high-dimensional cases.
660 A possible idea would be to look at techniques developed for three dimensional problems in the finite element community. The use of criteria I and III is probably here more straightforward.

The field of multi-element polynomial chaos approaches would benefit from further work in the partitioning techniques. Instead of splitting cells in two
665 equal parts, one may think of a cleverer way to find a not-arbitrary splitting. Furthermore, during long-term integration, it is possible that some areas in the stochastic space which need refinement at a particular time because becoming difficult to approximate polynomially, start to be smooth again further in time. It means that, in the concept of adaptive meshing, the possibility of joining
670 elements and obtaining a coarser mesh locally would be beneficial for the overall computing time. The meshing methodology presented here unfortunately only considers division but not junction of elements. A possible lead to follow would be to analyse if the neighbouring approximations are close enough to the current element approximation, which would imply some kind of local smoothness and
675 the possibility of an eventual locally coarser mesh. Further study on this aspect of adaptive and optimal meshing will be considered by the authors.

In subsequent papers, the proposed MEgPC methodologies will be applied to the problem of uncertainty quantification in acoustic pumping. In systems driven by experimental acoustic measurements, uncertainties in modelling pa-
680 rameters are common and a polynomial chaos approach is an efficient tool to accurately predict the response dispersion. The existence of a nonlinear energy sink used in such application also supports the need of stochastic space partitioning.

Acknowledgment

685 The authors gratefully acknowledge the support for this work provided by the Mechanics and Complexity Laboratory of Excellence (MEC LabEx). The

authors would also like to thank colleagues Prof. B. Cochelin for support and advice and T. Le Gouic for his insights into Wasserstein distances.

References

- 690 [1] G. Schuëller, Computation stochastic mechanics - recent advances, Computers & Structures 79 (2001) 2225–2234. [doi:10.1016/S0045-7949\(01\)00078-5](https://doi.org/10.1016/S0045-7949(01)00078-5).
- [2] D. Xiu, Fast numerical methods for stochastic computation: a review, Communications in Computational Physics 5 (2009) 242–272. [doi:10.1.1.148.5499](https://doi.org/10.1.1.148.5499).
- 695 [3] A. Nouy, Recent developments in spectral stochastic methods for the numerical solution of stochastic partial differential equations, Archives of Computational Methods in Engineering 16 (2009) 251–285. [doi:10.1007/s11831-009-9034-5](https://doi.org/10.1007/s11831-009-9034-5).
- 700 [4] G. Fishman, Monte Carlo: Concepts, Algorithms, and Applications, New York, 1996. [doi:10.1007/978-1-4757-2553-7](https://doi.org/10.1007/978-1-4757-2553-7).
- [5] W.-L. Loh, On latin hypercube sampling, The Annals of Statistics 24 (1996) 2058–2080. [doi:10.1214/aos/1069362310](https://doi.org/10.1214/aos/1069362310).
- [6] I. Sobol, On quasi-monte carlo integrations, Mathematics and Computers in Simulation 47 (1988) 103–112. [doi:10.1016/S0378-4754\(98\)00096-2](https://doi.org/10.1016/S0378-4754(98)00096-2).
- 705 [7] J. Spall, An overview of the simultaneous perturbation method for efficient optimization, Johns Hopkins APL Technical Digest 19 (1998) 482–492.
- [8] N. Wiener, The homogeneous chaos, American Journal of Mathematics 60 (1938) 897–936. [doi:10.2307/2371268](https://doi.org/10.2307/2371268).
- 710 [9] R. Ghanem, P. D. Spanos, Stochastic Finite Elements: A Spectral Approach, Springer-Verlag, 1991. [doi:10.1007/978-1-4612-3094-6](https://doi.org/10.1007/978-1-4612-3094-6).
- [10] I. Babuška, F. Nobile, R. Tempone, A stochastic collocation method for elliptic partial differential equations with random input data, SIAM Review 52 (2010) 317–355. [doi:10.1137/100786356](https://doi.org/10.1137/100786356).
- 715 [11] P. Mohan, P. Nair, A. Keane, Multi-element stochastic reduced basis methods, Computer Methods in Applied Mechanics and Engineering 197 (2008) 1495–1506. [doi:10.1016/j.cma.2007.11.021](https://doi.org/10.1016/j.cma.2007.11.021).
- [12] J. Foo, X. Wan, G. Karniadakis, The multi-element probabilistic collocation method (ME-PCM): Error analysis and applications, Journal of Computational Physics 227 (2008) 9572–9595. [doi:10.1016/j.jcp.2008.07.009](https://doi.org/10.1016/j.jcp.2008.07.009).
- 720

- [13] J. Foo, G. Karniadakis, Multi-element probabilistic collocation method in high dimensions, *Journal of Computational Physics* 229 (2010) 1536–1557. doi:10.1016/j.jcp.2009.10.043.
- 725 [14] X. Ma, N. Zabaras, An adaptive hierarchical sparse grid collocation algorithm for the solution of stochastic differential equations, *Journal of Computational Physics* 228 (2009) 3084–3113. doi:10.1016/j.jcp.2009.01.006.
- 730 [15] J. Witteveen, G. Iaccarino, Subcell resolution in simple stochastic collocation for spatial discontinuities, *Journal of Computational Physics* 251 (2013) 17–52. doi:10.1016/j.jcp.2013.05.035.
- [16] D. Xiu, G. Karniadakis, The wiener-askey polynomial chaos for stochastic differential equations, *SIAM Journal on Scientific Computing* 24 (2002) 619–644. doi:10.1137/S1064827501387826.
- 735 [17] C. Soize, R. Ghanem, Physical systems with random uncertainties: Chaos representations with arbitrary probability measure, *SIAM Journal on Scientific Computing* 26 (2) (2004) 395–410. doi:10.1137/S1064827503424505.
- [18] X. Wan, G. Karniadakis, Multi-element generalized polynomial chaos for arbitrary probability measures, *SIAM Journal on Scientific Computing* 28 (2006) 901–928. doi:10.1137/050627630.
- 740 [19] X. Wan, G. Karniadakis, An adaptive multi-element generalized polynomial chaos method for stochastic differential equations, *Journal of Computational Physics* 209 (2005) 617–642. doi:10.1016/j.jcp.2005.03.023.
- [20] M. Gerritsma, J. Van Der Steen, P. Vos, G. Karniadakis, Time-dependent generalized polynomial chaos, *Journal of Computational Physics* 229 (2010) 8333–8363. doi:10.1016/j.jcp.2010.07.020.
- 745 [21] G. Kewlani, K. Iagnemma, A multi-element generalized polynomial chaos approach to analysis of mobile robot dynamics under uncertainty, in: *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, St Louis, USA, 2009, pp. 1177–82. doi:10.1109/IROS.2009.5354420.
- 750 [22] J. Le Meitour, D. Lucor, J.-C. Chassaing, Prediction of stochastic limit cycle oscillations using an adaptive polynomial chaos method, *Journal of Aeroelasticity and Structural Dynamics* 2 (2010) 3–22. doi:10.3293/asdj.2010.4.
- 755 [23] D. Xiu, *Numerical Methods for Stochastic Computations – A Spectral Method Approach*, Princeton University Press, 2010.
- [24] R. Cameron, W. Martin, The orthogonal development of non-linear functionals in series of Fourier-Hermite functionals, *Annals of Mathematics* 48 (1947) 385–392. doi:10.2307/1969178.
- 760

- [25] W. Press, S. Teukolsky, W. Vetterling, B. Flannery, Numerical recipes : The art of scientific computing, 3rd Edition, Cambridge University Press, 2007.
- 765 [26] E. Sarrouy, O. Dessombz, J.-J. Sinou, Piecewise polynomial chaos expansion with an application to brake squeal of a linear brake system, Journal of Sound and Vibration 332 (2013) 577–594. [doi:10.1016/j.jsv.2012.09.009](https://doi.org/10.1016/j.jsv.2012.09.009).
- [27] P. Ladevèze, J.-P. Pelle, La maîtrise du calcul en mécanique linéaire et non-linéaire, Hermes Science Publication, 2001.
- 770 [28] O. C. Zienkiewicz, J. Z. Zhu, A simple error estimator and adaptive procedure for practical engineering analysis, International Journal for Numerical Methods in Engineering 24 (1987) 337–357. [doi:10.1002/nme.1620240206](https://doi.org/10.1002/nme.1620240206).
- 775 [29] O. C. Zienkiewicz, J. Z. Zhu, The superconvergent patch recovery and a posteriori error estimates. part I: the recovery technique., International Journal for Numerical Methods in Engineering 33 (1992) 1331–1364. [doi:10.1002/nme.1620330702](https://doi.org/10.1002/nme.1620330702).
- [30] O. C. Zienkiewicz, J. Z. Zhu, The superconvergent patch recovery and a posteriori error estimates. part II: error estimates and adaptivity., International Journal for Numerical Methods in Engineering 33 (1992) 1365–1382. [doi:10.1002/nme.1620330703](https://doi.org/10.1002/nme.1620330703).
- 780 [31] C. Villani, Optimal transport. Old and new, Vol. 338 of Grundlehren der mathematischen Wissenschaften, Springer-Verlag, Berlin, 2009. [doi:10.1007/978-3-540-71050-9](https://doi.org/10.1007/978-3-540-71050-9).
- 785 [32] E. del Barrio, J. A. Cuesta-Albertos, C. Matrn, J. M. Rodríguez-Rodríguez, Tests of goodness of fit based on the L_2 -Wasserstein distance, The Annals of Statistics 27 (4) (1999) 1230–1239. [doi:10.1214/aos/1017938923](https://doi.org/10.1214/aos/1017938923).
- [33] E. Sarrouy, J.-J. Sinou, Advances in Vibration Analysis Research, InTech, 2011, Ch. Non-Linear Periodic and Quasi-Periodic Vibrations in Mechanical Systems - On the use of the Harmonic Balance Methods, pp. 419–434. [doi:10.5772/15638](https://doi.org/10.5772/15638).
- 790 [34] S. Orszag, L. Bissonnette, Dynamical properties of truncated Wiener-Hermite expansions., Physics of Fluids 10 (1967) 2603–2613. [doi:10.1063/1.1762082](https://doi.org/10.1063/1.1762082).
- 795 [35] N. Agarwal, N. Aluru, A domain stochastic collocation approach for analysis of mems under uncertainties, Journal of Computational Physics 228 (2009) 7662–7688. [doi:10.1016/j.jcp.2009.07.014](https://doi.org/10.1016/j.jcp.2009.07.014).

Appendix A. Legendre polynomials

Legendre polynomials can be defined using Bonnet's recursion formula:

$$\begin{cases} \psi_0(x) = 1, \psi_1(x) = x \\ \forall k \geq 1, (k+1)\psi_{k+1}(x) = (2k+1)x\psi_k(x) - k\psi_{k-1}(x) \end{cases} \quad (\text{A.1})$$

800 The first 4 polynomials are:

$$\begin{aligned} \psi_0(x) &= 1 & \psi_2(x) &= \frac{1}{2}(3x^2 - 1) \\ \psi_1(x) &= x & \psi_3(x) &= \frac{1}{2}(5x^3 - 3x) \end{aligned} \quad (\text{A.2})$$

This set of polynomials is orthogonal with respect to the following scalar product:

$$\langle f, g \rangle = \frac{1}{2} \int_{-1}^1 f(x)g(x)dx \quad (\text{A.3})$$

The weight function $w(x) = 1/2$ is then equal to ξ pdf. Subsequently, one has:

$$\langle \psi_i, \psi_j \rangle = \mathbb{E}[\psi_i(\xi)\psi_j(\xi)] \quad (\text{A.4})$$

In particular, it is worth noticing the following results:

$$\langle \psi_0, \psi_0 \rangle = \mathbb{E}[\psi_0(\xi)\psi_0(\xi)] = 1, \quad (\text{A.5})$$

805

$$\forall k \geq 1, \langle \psi_k, \psi_0 \rangle = \mathbb{E}[\psi_k(\xi)\psi_0(\xi)] = 0. \quad (\text{A.6})$$

In the frame of a polynomial chaos expansion with Eq. (3) notations, this implies that the mean and variance are given by:

$$\mu_U = \mathbb{E}[U] = u_0 \quad (\text{A.7})$$

$$\sigma_U^2 = \mathbb{E}[(U - \mathbb{E}[U])^2] = \sum_{k=1}^{N_p} u_k \langle \psi_k, \psi_k \rangle \quad (\text{A.8})$$

Appendix B. Detailed expression for criterion II local error

810

The local error for criterion II over the n -th element can be decomposed into a weight, β_n , relative to the element size, and an error computed in the local basis η_n^{II} , leading to $\epsilon_n^{\text{II}} = \beta_n \eta_n^{\text{II}}$. η_n^{II} is calculated from Eq. (21) and makes use of the value of \tilde{g} in ξ_n , which is obtained for non-extremal breaks using linear interpolation between $\tilde{g}(\alpha_{n-1})$ and $\tilde{g}(\alpha_n)$:

$$\tilde{g}(\xi_n) = \frac{\beta_n \hat{g}_{n-1} + \beta_{n-1} \hat{g}_n}{\beta_n + \beta_{n-1}}, \quad n \in \llbracket 2, N_e \rrbracket. \quad (\text{B.1})$$

815 For non extremal elements (i.e. $n \in \llbracket 2, N_e - 1 \rrbracket$) η_n^{II} becomes:

$$\eta_n^{\text{II}} = \frac{\beta_n}{\sum_{n=1}^{N_e} \sigma_{U_n}} \left(\frac{|\hat{g}_{n-1} - \hat{g}_n|}{\beta_n + \beta_{n-1}} + \frac{|\hat{g}_{n+1} - \hat{g}_n|}{\beta_n + \beta_{n+1}} \right). \quad (\text{B.2})$$

For extremal elements the formula depends on the positivity of the extrapolated values of \tilde{g} . This positivity does not have to be tested and the following formulas handle all the cases:

$$\eta_1^{\text{II}} = \frac{1}{\sum_{n=1}^{N_e} \sigma_{U_n}} \left(\frac{\beta_1 |\hat{g}_2 - \hat{g}_1|}{\beta_1 + \beta_2} + \min \left(\frac{\beta_1 |\hat{g}_2 - \hat{g}_1|}{\beta_1 + \beta_2}, \hat{g}_1 \right) \right), \quad (\text{B.3})$$

and

$$\eta_{N_e}^{\text{II}} = \frac{1}{\sum_{n=1}^{N_e} \sigma_{U_n}} \left(\frac{\beta_{N_e} |\hat{g}_{N_e} - \hat{g}_{N_e-1}|}{\beta_{N_e} + \beta_{N_e-1}} + \min \left(\frac{\beta_{N_e} |\hat{g}_{N_e} - \hat{g}_{N_e-1}|}{\beta_{N_e} + \beta_{N_e-1}}, \hat{g}_{N_e} \right) \right). \quad (\text{B.4})$$

820 Appendix C. Residue expression for criterion III in the case of ODEs

When the set of governing equation is an ODE, the residue calculation depends on the method used to evaluate the PC expansion coefficients.

If an intrusive method is used, the initial set of equations $\dot{u} = f(u, t)$ is turned into a larger system which usually depends on the element I_n and can be written as:

$$\dot{U}_n^{\text{PC}} = F_n(U_n^{\text{PC}}, t), \quad (\text{C.1})$$

where U_n^{PC} gathers the PC expansion coefficients $\tilde{u}_{k,n}$.

Having computed U_n^{PC} for a given time t enables to compute its derivative with respect to time using the F_n function. Realisations of \dot{U}_n can then be evaluated using a formula similar to Eq. (7) and the local residue can be defined over each element I_n as the difference between \dot{U}_n and $f(U_n, t)$.

830 If a non-intrusive method is used the PC decomposition of \dot{U}_n can be evaluated the same way as U_n decomposition. Once realisations of U_n are calculated at collocation points, corresponding \dot{U}_n values are computed using f . Applying the same collocation method for \dot{U}_n as for U_n then provides its PC expansion
835 over the n -th element. The residue is finally defined in the same way as in the intrusive case, that is, as the difference between \dot{U}_n and $f(U_n, t)$.

Appendix D. Illustration of Wasserstein distance

The Wasserstein distance is illustrated in this appendix via several examples. In what follows, a target random variable U and its approximation \hat{U}
840 are compared considering different error indicators : the relative error in mean ϵ_E (defined in Eq. (D.1)), the relative error in variance ϵ_σ (also defined in

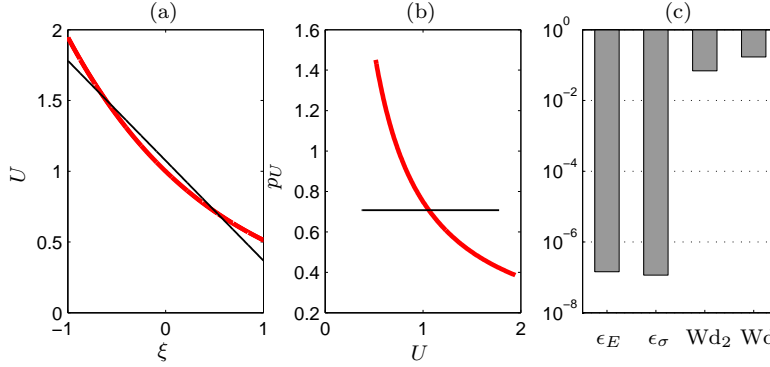


Figure D.13: Wasserstein distance illustration: Advantages, Example 1. Thick (red) line: target random variable; thin (black) line: approximated random variable.

Eq. (D.1)), the Wasserstein distance Wd_2 (defined in Eq. (30)), and its dimensionless version Wd (defined in Eq. (31)) that what used throughout Section 4:

$$\epsilon_E = \frac{|\mu_U - \mu_{\hat{U}}|}{|\mu_U|} \text{ and } \epsilon_\sigma = \frac{|\sigma_U - \sigma_{\hat{U}}|}{|\sigma_U|} \quad (D.1)$$

845 where $\mu_X = E[X]$ and $\sigma_X = \sqrt{\text{Var}[X]}$.

The first example shows the advantage of the Wasserstein distance when one is interested not only in the first two moments but also in the global distribution. Here, the approximation \hat{U} is chosen such that it has the same mean and variance as the target random variable U .

- 850 • Example 1: the target random variable is $U = 1 + e^{-\xi/s}$ with $s = 1.5$ and the approximation is $\hat{U} = a + b\sqrt{3}\xi$ with $a = E[U] = s \sinh(1/s)$ and $b = \sqrt{\text{Var}[U]} = \sqrt{s \sinh(1/s) (\cosh(1/s) - s \sinh(1/s))}$. Results are displayed in Fig. D.13.

In this first example, one can see from Fig. D.13 that \hat{U} is a rather bad approximation of U (pane (a)) and that the resulting pdf is also very different (pane (b)) from the target one. However, relative errors ϵ_E and ϵ_σ are both very small ($\approx 10^{-7}$). This was expected by construction of \hat{U} . This said, it is to be noted that the Wasserstein distances Wd_2 and Wd are not “small” ($\approx 10^{-1}$) and then, they seem able to provide more interesting information considering the inaccuracy of the pdf approximation.

860 As a “small” number is a very relative notion, the next three examples try to quantify it. For each example, the target random variable U is approximated by polynomials with increasing degrees. These polynomials are computed using a simple intrusive Polynomial Chaos expansion. For each approximation, the pdf is built, the 4 error indicators are evaluated and their evolution is monitored.

- Example 2: The reference random variable is $U = 1 + \cos(\pi \xi) + 2 \sin(\pi \xi)$; polynomials with degree 2, 4 and 6 are used for \hat{U} approximation. Results are illustrated in Fig. D.14.

- Example 3 (mechanical system response near resonance): The reference random variable is $U = 10/|\omega_0(\xi)^2 - \omega^2 + 2j \times 0.03 \times \omega_0(\xi)\omega|$ where $\omega_0(\xi) = 1 + 0.2\xi$, $\omega = 0.90$. This depicts the amplitude of a simple mass-spring-damper oscillator similar to the one studied in Section 4.2.1. Here, the target random variable is not the cosine component of the response but its amplitude. Polynomials with degree 20, 35 and 50 are used for \hat{U} approximation. Results are displayed in Fig. D.15.
- Example 4: The reference random variable is $U = e^{-4\xi^2}$. Polynomials with degree 6, 10 and 14 are used for \hat{U} approximation. Results are shown in Fig. D.16.

In each case, the lowest polynomial degree simulated returns a very bad approximation considering the pdf while the highest degree returns a rather good one.

As Wd is a dimensionless, or relative, measure, it is expected not to vary too much with respect to the order of magnitude of U , and therefore its quantitative value is more meaningful and consistent from one example to another than Wd_2 value. Moreover, from this set of illustrations, it can be said that $Wd < 10^{-2}$ seems an acceptable condition to get the skeleton of the pdf, $Wd < 10^{-3}$ looks like a good criterion to get an accurate description of the pdf and $Wd < 10^{-4}$ seems required to get a perfect match.

The quantitative conclusions that can be drawn from these examples are not proved in any way and the proposed Wd limits may not be adequate for every possible case. However, this collection of very different test cases still provides some useful information on the meaning of the order of magnitude of Wd.

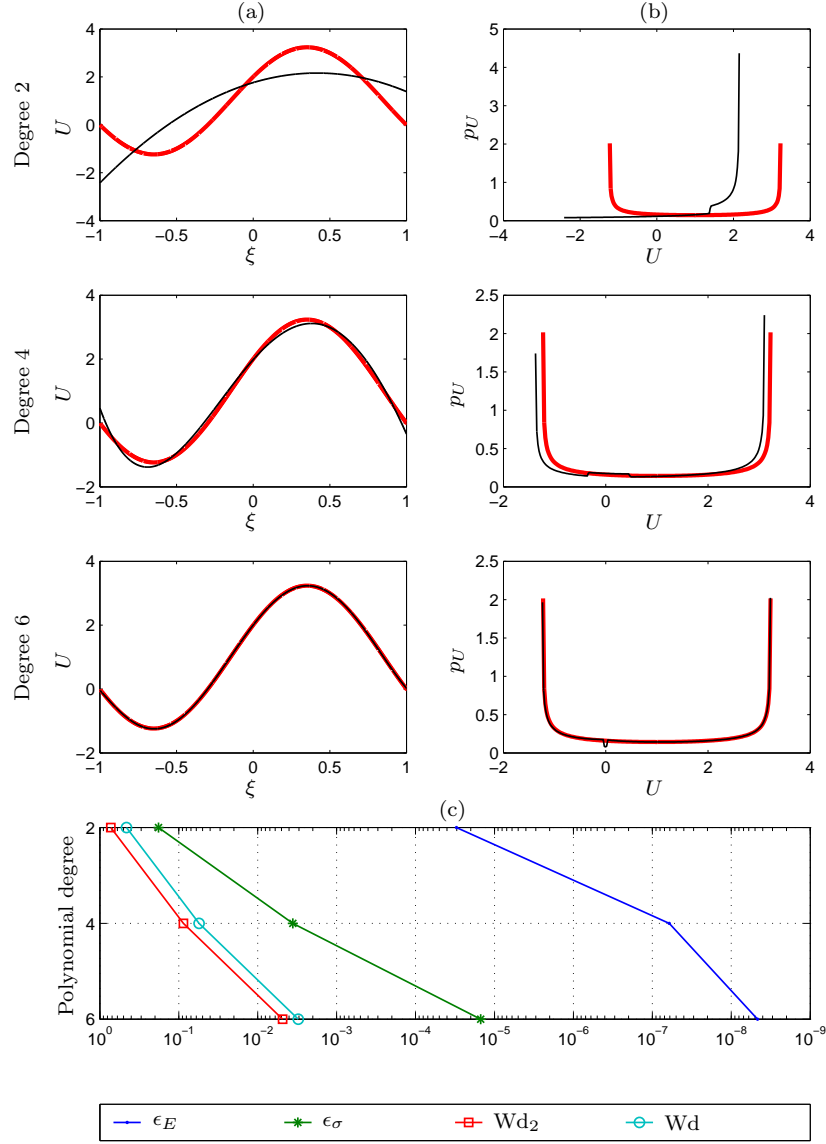


Figure D.14: Wasserstein distance illustration: Value, Example 2. Thick (red) line: target random variable; thin (black) line: approximated random variable.

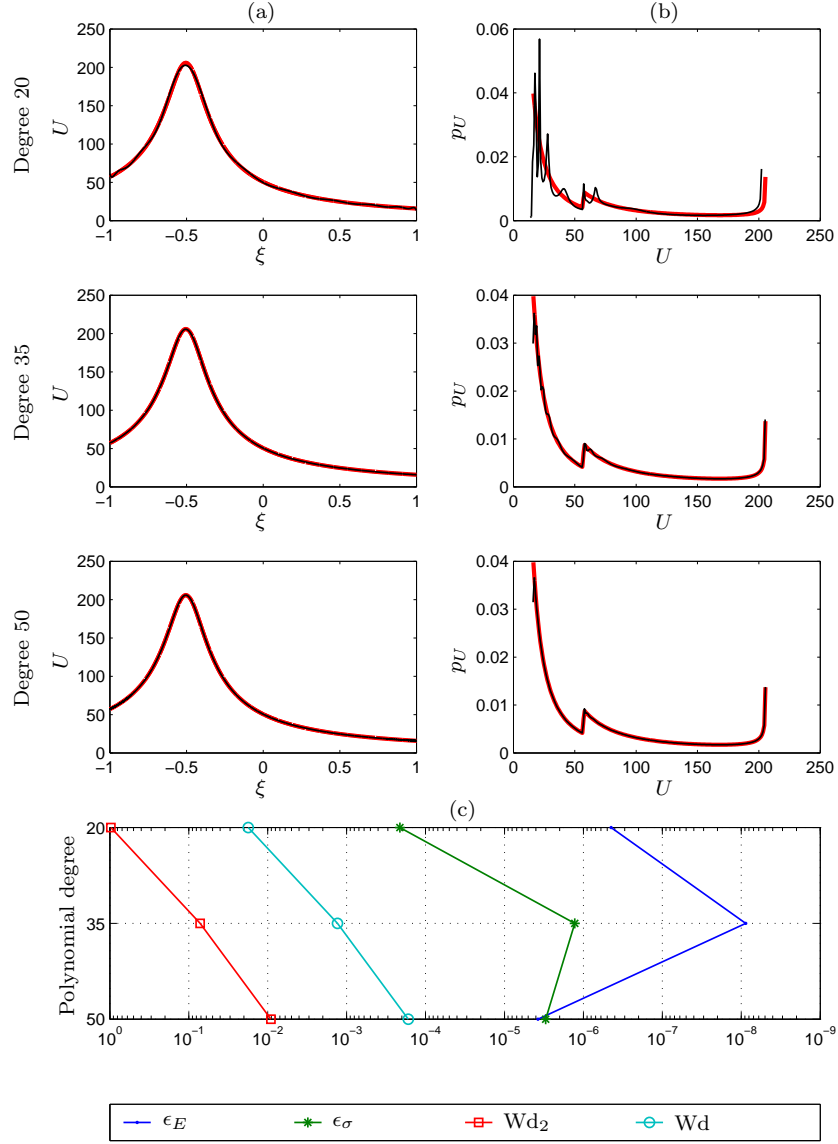


Figure D.15: Wasserstein distance illustration: Value, Example 3. Thick (red) line: target random variable; thin (black) line: approximated random variable.

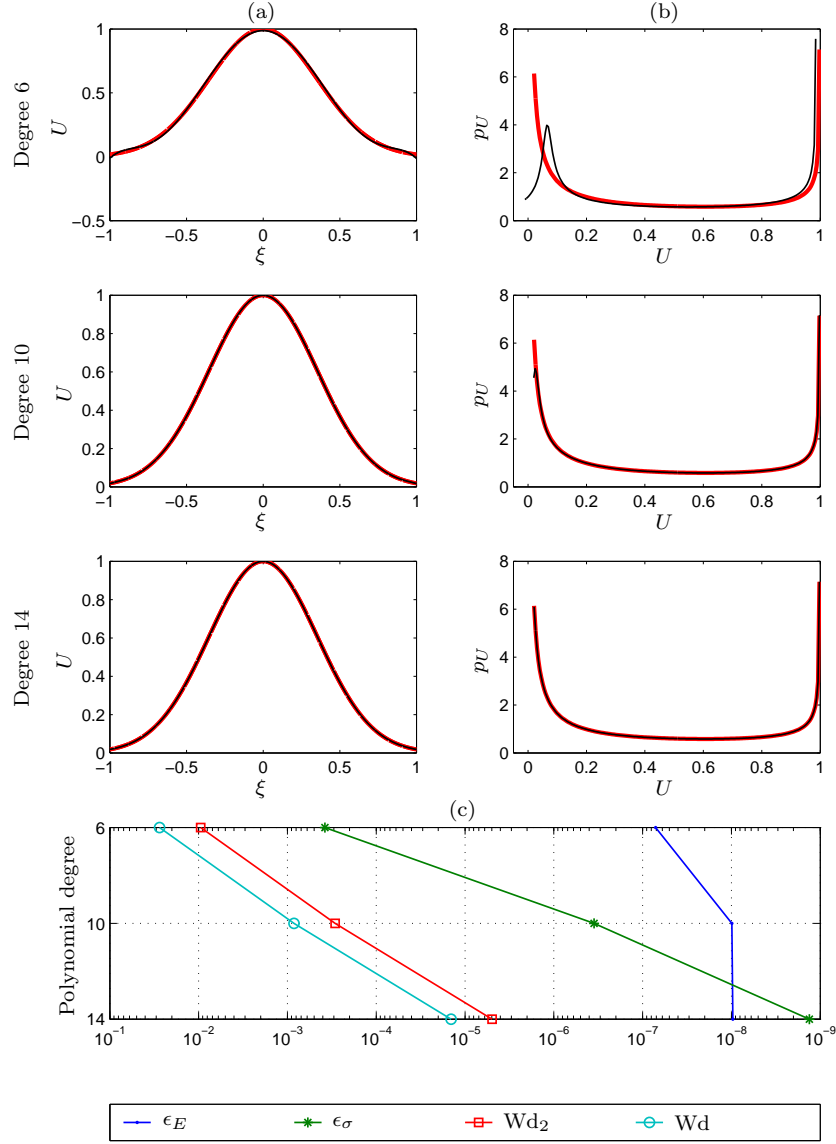


Figure D.16: Wasserstein distance illustration: Value, Example 4. Thick (red) line: target random variable; thin (black) line: approximated random variable.